**CRS report for Congress(Data Mining and Homeland Security)**

We used data mining for fraud detecting, risk assessing and retailing of product.it also involve data mining tools to discover unknown, valid pattern and relationship in large data records. The identification and tracking of activities such as money transfers, travel record, immigration record and communications. Some observers suggest that data mining should be used as a means to identify terrorist activities, such as money transfers and communications, and to identify and track individual terrorists themselves, such as through travel and immigration records. There are so many companies collecting customer data and develop models that predict whether a customer is a good credit risk etc.

We have some limitation in data mining such as data mining does not tell the user the value or significance of these patterns. These types of determinations must be made by the user. A second limitation is that while data mining can identify connections between behaviors and/or variables, it does not necessarily identify a causal relationship.

This paper is also focus on some issue related to data mining .for example accuracy and completeness of the data being analyzed. A second issue is the interoperability of the data mining software and databases being used by different agencies. A third issue is mission creep, or the use of data for purposes other than for which the data were originally collected. And fourth issue is privacy. Following are the term refer in data mining issue.

Interoperability refers to the ability of a computer system and/or data to work with other systems or data using common standards or processes. Mission creep refers to the use of data for purposes other than that for which the data was originally collected. Some observers contend that tradeoffs may need to be made regarding privacy to ensure security. Other observers suggest that existing laws and regulations regarding privacy protections are adequate, and that these initiatives do not pose any threats to privacy.

In the later part some limitation as a terrorist detection tool is discussed.1st data mining required skilled technical and analytical specialists who can structure the analysis and interpret the output.

2nd although data mining can help reveal patterns and relationships, it does not tell user the value or significance of these patterns.

Data quality refers to the accuracy and completeness of the data. Data quality can also be affected by the structure and consistency of the data being analyzed. To improve data quality, it is sometimes necessary to “clean” the data, which can involve the removal of duplicate records, normalizing the values used to represent information in the database, accounting for missing data points, removing unneeded data fields etc.

In order to implement current paper approach ,they have developed a solution name as Factual Analysis Criminal Threat Solution (FACTS) which was described as a “technological, investigative tool allowing query-based searches of available state and public records in the data reference repository.it allowed an authorized user to search “dynamically combined records from disparate datasets” based on partial information, and will “assemble” the results.

The data reference repository purportedly excluded data such as telemarketing call lists, direct mail mailing lists, airline reservations or travel records, frequent flyer/hotel stay program

membership or activity, magazine subscriptions, information about purchases made at retailers or over the Internet, telephone calling logs or records, credit or debit card numbers, mortgage or car payment information, bank account numbers or balance information, birth certificates, marriage licenses, divorce decrees, or utility bill payment information.

In the last part there is a discussion of legislation on user’s information collection which does not disturb the privacy and security of a person.